COST VS CREDIBILITY: HOW MUCH V&V IS ENOUGH?

Dr. Paul R. Muessig
Naval Air Warfare Center, Weapons Division
Code 418100D (SMART Project Office)
China Lake, CA 93555-6001

ABSTRACT

This paper addresses the perennial issue, "How much V&V is enough for an accreditation?" or, perhaps more to the point, "How much do you have to spend to convince someone that your model is any good?". The answer is discussed in terms of our experience in providing accreditation support for weapons system programs within the Department of Defense (DoD).

We begin by describing an integrated V&V process developed by the SMART Project, both in overview and in detailed Work Breakdown Structure (WBS) form. We then describe how V&V tasking over the life of the project was tracked and related to WBS elements, and how average costs of each V&V task and product were developed. A table showing levels of effort (LOEs) and fully loaded costs for each phase of V&V is also presented.

The V&V process is then embedded within the context of a five-step accreditation support mechanism, whose focus is on the development of application specific M&S acceptance criteria prior to setting V&V requirements. By relating V&V requirements to objective acceptance criteria, and by structuring the accreditation plan around these requirements, precious V&V resources are not wasted chasing unneeded V&V products.

The paper concludes by demonstrating that the cost of establishing M&S credibility is governed by two factors: (1) development of unambiguous M&S acceptance criteria prior to V&V, and (2) the availability of prior V&V results. Absent either of these two factors the cost of accreditation must rise substantially, either due to overscoping of the V&V effort, or to lack of synergism between current and prior accreditation efforts, or both.

INTRODUCTION

The recent focus on the requirement for M&S credibility has come to be balanced by an equal concern for the cost of the V&V activities that contribute to it. The lack of a coherent scheme tying V&V products to M&S credibility requirements has hindered the identification of a minimum set of V&V activities that meet those requirements. The result has been a tendency either to overestimate V&V requirements (resulting in the frequently overheard lament that "V&V is too expensive"), or to dilute them (out of frustration) to the point of insignificance (resulting in accreditation by fiat).

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
00Aug2001	N/A	-

Title and Subtitle Cost VS Credibility: How Much V&V Is Enough?	**Contract Number**
	Grant Number
	Program Element Number
Author(s) Paul Muessig	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Naval Air Warfare Center, Weapons Division (Code 418100D) China Lake, California 93555-6001	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes

Abstract
see report

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** SAR

Number of Pages
13

Contributing to the inability to quantify the cost of V&V has been a lack of generically applicable cost data from efforts with broad experience in V&V tasking. V&V cost data are mostly anecdotal, and derived from efforts that may not have sampled widely enough from the V&V menu. Moreover, existing cost data are usually unrelatable to well defined V&V products that can be tied to specific M&S credibility requirements. With the maturation and stabilization of V&V techniques, however, has come an ability to relate V&V costs to M&S credibility in a more quantifiable way. The SMART Project has contributed to this quantification in three ways:

1. From the myriad of possible V&V tasks and products, it has identified those which contribute most heavily to M&S accreditation decisions, and has developed significant practical expertise in each of them;

2. From this core list of V&V tasks and products, it has developed an integrated V&V process (including elements of configuration management, or C/M) that generates the most frequently required V&V products at various levels of detail and cost, and;

3. It has embedded both the V&V process and it's products within the context of accreditation decision making by developing (and applying) a procedure for defining objective requirements for M&S credibility based on intended M&S applications.

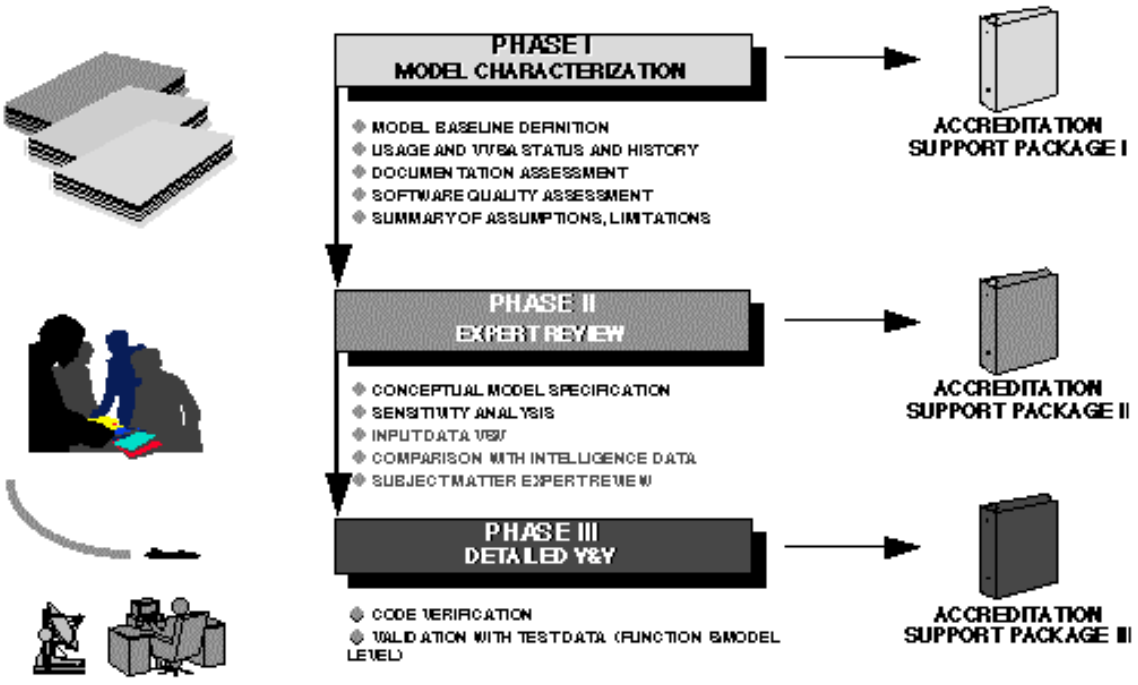

Figure 1: SMART's Three Phase V&V Process

SMART V&V PROCESS OVERVIEW

To determine what V&V information products were most essential to provide, SMART

conducted an Accreditation Requirements Study , which identified the types of V&V information currently being used to support accreditation decisions for military M&S, and embedded them within the context of three contributing factors: (1) the M&S management and VV&A policies, procedures and guidelines emerging across the services and DoD; (2) the accreditation support hierarchy developed by the MORS SIMVAL program, and; (3) the developing SMART V&V process. The result was the three phase V&V process depicted in figure 1. The three "phases" of V&V activity are model characterization, expert review, and detailed V&V. Although originally developed for application to engagement level air combat survivability simulations, the process is in no way dependent on this application.

V&V Phase Descriptions

Phase I V&V is geared toward characterizing a model. Typical questions addressed are: How is the model managed and supported? What has it been used for, by whom, and was it accredited for that use? What is its V&V history? How well is it documented? What is the quality of the software? What are the model's known assumptions, limitations and errors? The end result of Phase I V&V is confidence that the model so characterized and controlled will produce consistent results across a spectrum of users and applications, and that its predictions have been used by others with similar applications.

Phase II V&V is aimed at a subjective determination of model "reasonableness" by means of a review of a collection of objective data. The review is conducted by subject matter experts (SMEs) in areas relevant to the model's functions (for example, radar systems engineers for radar functions, guidance and control specialists for missile functions, etc.). The objective data that provide grist for the review mill consist of: input data verification and validation (coming to be known in the defense community as "VV&C", where the C stands for Certification to distinguish it from the A in M&S VV&A); comparison of model outputs with intelligence data or best estimates; sensitivity analysis results, and; a summary of known assumptions, limitations and errors. These objective data are reviewed in light of the intended application of the M&S to arrive at subjective determination that it is "good enough" for the purpose at hand. Although fraught with the pitfalls inherent in subjectivity, the Phase II expert review, in conjunction with Phase I model characterization, provides the best possible assessment of the adequacy of macro-level model results short of detailed V&V (Phase III).

Phase III V&V is "classical" V&V: the kind of V&V everyone is afraid of. It typically consists of line-by-line verification of the code, including desk checking, software testing and comparison to design specifications, coupled with extensive comparisons of model predictions with all available sources of test data at both the detailed M&S functional level and the overall M&S output level. Because of its cost, Phase III V&V activities are not (or at least, should not) be performed on an entire model without reference to an application requirement. Rather, the scope of Phase III V&V should be tailored to each specific application, and only those M&S functions essential for use in that application should be subjected to this rigorous level of activity.

All phases of V&V activity are incremental and interdependent, viz., Phase II accreditation support products build on those produced in Phase I, and so forth. All phases have well defined tasks and well defined products chosen to maximize applicability to typical accreditation

decisions. All phases result in a standardized Accreditation Support Package that facilitates accreditation of M&S at various levels of detail. These levels are characterized by how much information about the model is available at the end of each phase. Taken as whole, the V&V process depicted in figure 1 provides an incremental approach to assessing M&S credibility that can be tailored to the needs of individual applications.

Cost Effectiveness of V&V Process

There are several aspects of this approach to V&V that have a potential to reduce the cost of M&S accreditation. First, low cost, low risk, high value V&V tasks are performed early in the process. These tasks are aimed at identifying M&S strengths and weaknesses in a short period of time, and permit an early assessment of M&S applicability to a wide variety of possible applications. Second, placing high cost, high risk V&V tasks at the end of the process allows time for development of objective criteria by which to minimize the scope of the Phase III effort. This avoids artificially inflated V&V cost estimates (and worries) and mitigates against the tendency to "V&V for it's own sake." It also guarantees that any detailed V&V performed will have maximal applicability to the accreditation decision. Third, the welldefined phases of V&V activity and the standard V&V reporting formats are essential to building cost effectiveness into accreditation decisions. V&V information developed in support of one accreditation benefits future users of the same model, who can build upon prior V&V information to reduce the scope of their own accreditation effort. Moreover, cost benefits for future users are guaranteed when each V&V effort is documented in a standard way, summarized in a central repository, and archived by a central accreditation support activity. Using this cycle, V&V becomes "market driven" by actual accreditation requirements, and no one sponsor has to pay for the V&V of an entire model.

COST ANALYSIS

But how much does all this really cost? To answer that question, we developed a Work Breakdown Structure (WBS) for each phase of V&V activity, and determined the level of effort (LOE) applied to each WBS element over the life of the project (three years to date). The result was a model-specific average LOE for each phase. Since these averages did not reflect completed products (none of the Phases were complete for any model at the time the data were collected), each LOE was normalized to the amount of progress actually achieved in each Phase. The normalized model-specific LOEs were then averaged to arrive at an estimated LOE for each phase that would apply to any model.

To convert LOE figures into cost figures, we calculated an equivalent hourly rate based on a weighted average of contractor LOE and individual labor costs over the life of the project, including other direct costs such as travel and documentation. This weighted average labor rate turned out to be approximately $87.50 per hour, or $14K per man-month, assuming a 160 hour man-month. We then multiplied the average LOE for each phase (in man-months) by the average dollar cost per man-month to arrive at the average dollar cost of each phase of V&V.

The results of this cost analysis are shown in figures 2 through 4 and table 1. Figures 2 through 4 show the detailed WBS and task sequencing for each phase. The first number in each WBS

element identifies its phase, the second number identifies its primary activity (1=verification, 2=validation, 3=configuration management), and the third number identifies the task sequence within a primary activity. Table 1 shows the average level of effort and dollar costs for each phase of V&V activity. The term "FE" used in the table refers to the "Functional Elements" that comprise the model. Identification of these FEs occurs in Phase I task 1.2.1 (Decompose Model). WBS tasks not listed in table 1 are those for which no reliable cost estimates were derivable because they varied with each application.

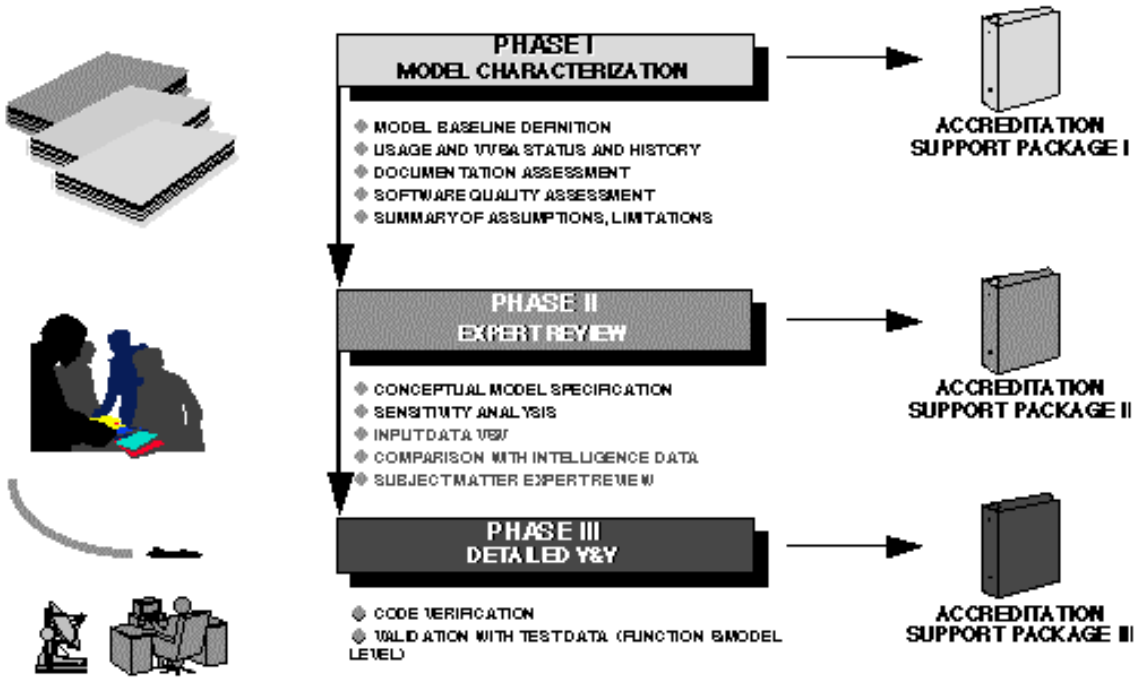

Figure 2: Phase I VV&CM Activities

Table 1 suggests some interesting characteristics about the incremental approach to V&V that should comfort those concerned about runaway V&V costs. For example, Phase I V&V tasks, whose products relate to the credibility of the model as a whole, has a fixed cost (i.e., about $250K), while the V&V tasks in Phases II and III have a variable cost related to the number of specific M&S functions requiring credibility to support a particular application. Since the total cost of Phase II and III V&V for two FEs exceeds the cost of the entire Phase I V&V effort according to table 1, it is clear that V&V costs are dominated by the number of FEs required to support a particular application. This suggests that significant attention must be paid to defining the smallest possible set of model functions required to support the model's use for a given application. It is in the rational and objective determination of this minimum set that the true path to cost effective V&V lies.

There are, however, other prerequisites for cost effective V&V. Consider, for example, two users of the same model who wish to accredit it for different applications. Assume further that both users have decided to complete Phase I, and that both require Phase II V&V for the same 5 FEs.

If they both work independently, and have no knowledge of each other's efforts, they will both spend 462K according to the data in table 1, for a total of almost a million dollars of V&V. At the end of all that, each user will have completed the same set of V&V tasks. Several important questions arise: How did each user interpret the same V&V task? Based on the different possible interpretations of each V&V task, how different do the resultant V&V reports look? Based on how different the V&V reports might look, how can future users benefit from two reports that interpret V&V differently, that say different things about the same model, and that report those things in different ways? Could the two efforts have worked together to reduce the total V&V cost burden while still accomplishing their individual accreditation objectives? How can future users benefit from the work these two independent efforts have done? A few moments reflection on possible answers to these questions reveals some enlightening suggestions for making V&V more cost effective.

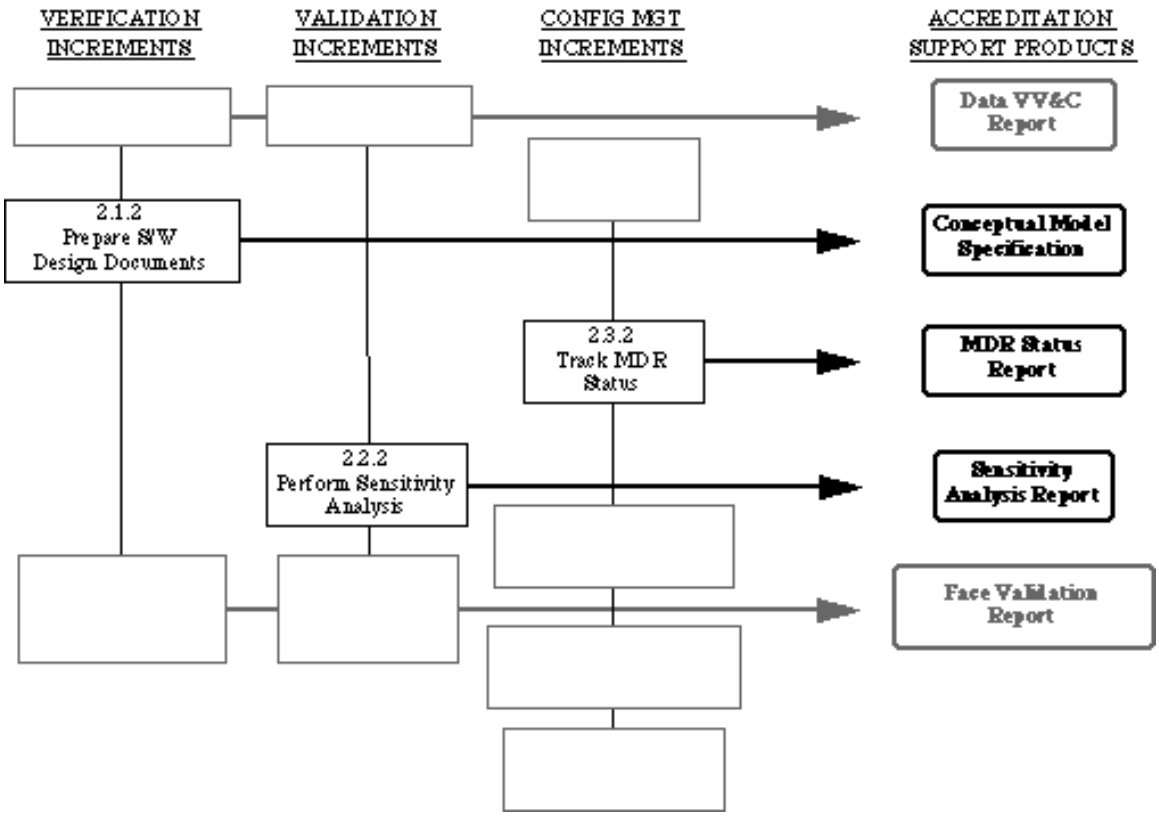

Figure: Phase II VV&CM Activities

Had each effort known of the other, a consolidated list of V&V tasks could have been drawn up from a coordinated review of M&S fidelity requirements for the two applications. In addition, a common understanding of V&V tasks could have been developed, as well as an understanding of how these tasks related to the previously defined fidelity requirements. Such a review could have led to a common V&V reporting format that would have served the needs of both applications without duplication of effort. By making this common V&V report available to other users of the

same model, V&V requirements in support of other applications could be reduced substantially, by forming the nucleus of a body of evidence supporting the model's credibility, a body to which other users with other applications could have contributed, in turn.

Three prerequisites for cost effective accreditation are suggested by the preceding discussion. Simply put, they are: (1) a standardized V&V process; (2) a standardized V&V reporting format, and; (3) a structured approach to developing objective M&S fidelity requirements based on application requirements. Lest the faint of heart (or the hard of head) take umbrage at the words "standardized" and "structured," let us hasten to explain that we do not construe these words to mean "mandated" and "inflexible." By standardized we mean that V&V techniques applicable to different classes of M&S should be well defined, understood and documented by the M&S community served by these techniques. Likewise, the product resulting from each V&V technique should be specified in such a way as to facilitate an orderly accumulation of the evidence for M&S credibility. Finally, an objective means of narrowing the scope of required V&V activities based on application requirements is necessary, to avoid the temptation to "do it all" or "make do with nothing."

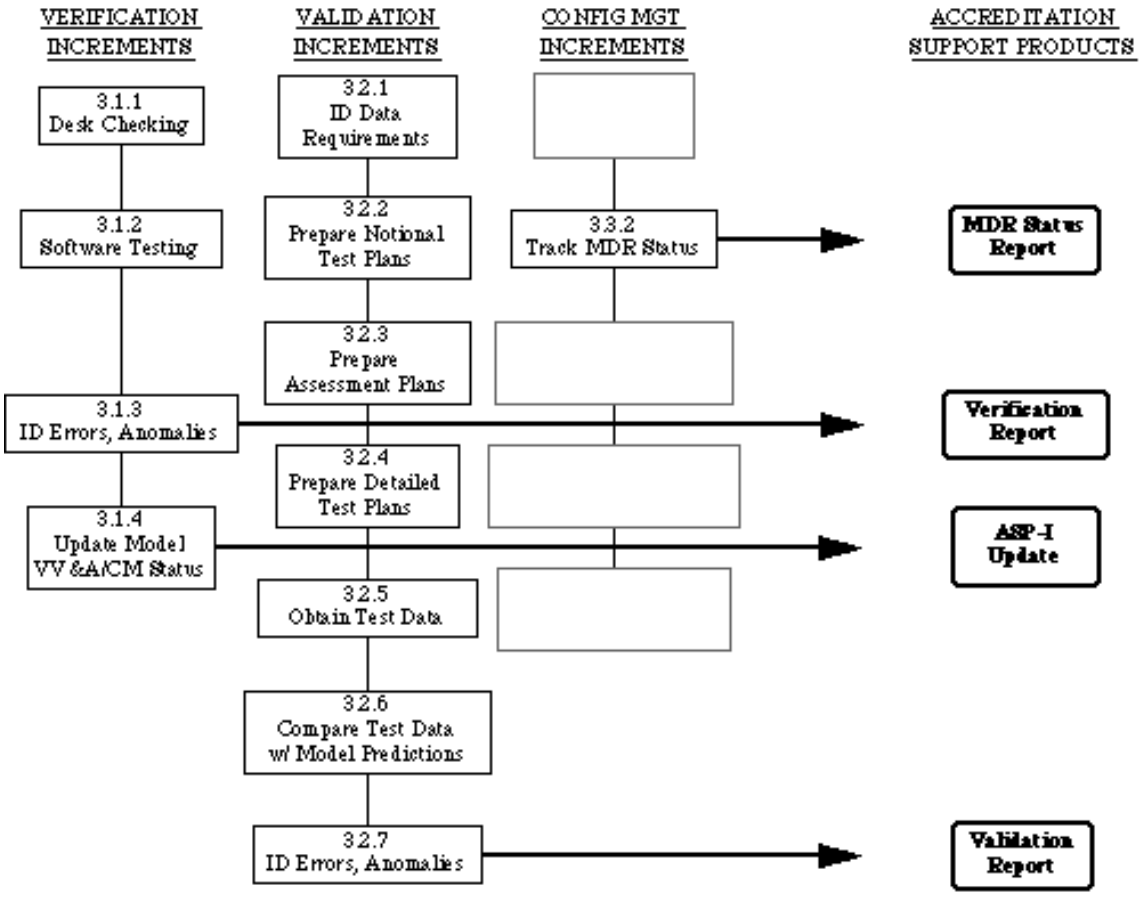

Figure 4: Phase III VV&CM Activities

The SMART Project has done all three of these things. It has developed and documented it's V&V process (applicable to the aircraft survivability M&S community) and made it available to that community for review and comment; it has developed standard V&V reporting formats and guidelines for use on aircraft survivability M&S, and is applying those guidelines to the to support the accreditation of aircraft survivability M&S, and; it has developed (and implemented) a process to identify objective fidelity requirements for aircraft survivability M&S based on case by case analysis of the applications those M&S are called upon to support. The first two items are available from the author for those who might wish to tailor these guidelines for their own applications or class of M&S. The last is shown in figure 5, and addresses the central theme of this paper.

WBS #	TASK NAME*	LOE (MM)	COST ($K)
1.1.1	Assess Current Documentation	3	
1.1.2	Assess Software Quality	3	
1.1.3	ID Assumptions, Limitations, Errors	3	
1.2.1	Decompose Model	2	
1.2.2	Define Functional Templates	1	
1.3.1	Define C/M Baseline	1	
1.3.2	Evaluate Existing C/M Procedures	2	
1.3.3	Survey Model History	3	
	PHASE I TOTALS	**18**	**252**
2.1.2	Prepare Software Design Documents	2/FE	
2.2.2	Perform Sensitivity Analysis	1/FE	
	PHASE II TOTALS	**3/FE**	**42/FE**
3.1.1-3.1.4	Code Verification Tasks	2/FE	
3.2.1-3.2.7	Validation Analysis Tasks	6/FE	
	PHASE III TOTALS	**8/FE**	**112/FE**
	GRAND TOTALS	**18 + 11/FE**	**252 + 154/FE**

Table 1: Typical V&V Costs

HOW MUCH V&V IS ENOUGH, ANYWAY?

As mentioned before, table 1 shows that V&V costs are driven by Phase II and III V&V requirements. Minimizing these requirements by identifying the key M&S functions necessary to support the application at hand is, therefore, essential to controlling V&V costs. Figure 5 depicts a generic five-step accreditation support process that focuses heavily on the determination of objective M&S acceptance criteria prior to execution of a V&V program.

The first step is an in-depth analysis of the application. Before any decisions are made about applying M&S to a given problem, the problem itself should be defined and articulated clearly enough to permit a precise specification of where M&S will play a role in the solution of the

problem and how they will contribute to the solution of that problem. An ill-defined or poorly understood problem is the most common reason for the failure to integrate M&S into analysis in a credible way. The problem analysis itself consists of four tasks: (1) a clear identification of study objectives; (2) development of an agreed-upon set of measures of effectiveness (MOEs) by which each study objective will be addressed and resolved; (3) an identification of which study objectives will be addressed or resolved using M&S, and: (4) an identification of the required predictive capabilities that M&S must have in order to be of value to the study MOEs (i.e., functional requirements). The mapping of study objectives, MOEs and M&S functional requirements is the single most important aspect of the accreditation support process because it forms a template for the entire M&S accreditation plan.

The next step in the accreditation support process is the development of M&S acceptance criteria. Having defined what "things" M&S will be required to do in the study, it remains to determine how well candidate M&S must do those "things." The answer to this question comes in the form of two types of acceptance criteria: operational requirements and fidelity requirements.

Operational requirements are "non-analytical" requirements, in that they do not contribute to evaluation of study MOEs directly. Instead, they specify hardware and software compatibility requirements (e.g., the M&S must run on a certain type of workstation under a certain operating system), pre- and post-processing requirements for M&S data (e.g., M&S inputs or outputs must be converted to special file formats), operations and training support requirements (e.g., candidate M&S cannot have license agreement or operator training requirements because there is no budget for such items, or no time for training), and so on.

Fidelity requirements are the hardest to define, and consist of a listing of how well each required M&S function must correlate to real world data in order for M&S outputs to be considered acceptable for the purpose at hand. This normally requires the development of a notional "error budget," whereby variations in M&S out-puts can be related to variations in study results via the MOEs. Fidelity requirements can be specified in terms of the standardized V&V information elements developed by SMART for incorporation into its ASP format. Although these elements provide a convenient checklist for fidelity requirements, however, the amount of V&V required to establish credibility for a particular application will still be dependent on a clear definition of the study problem. It is now clear why a precise relationship between the study problem, its MOEs and M&S functional requirements (developed under the first task) is essential.

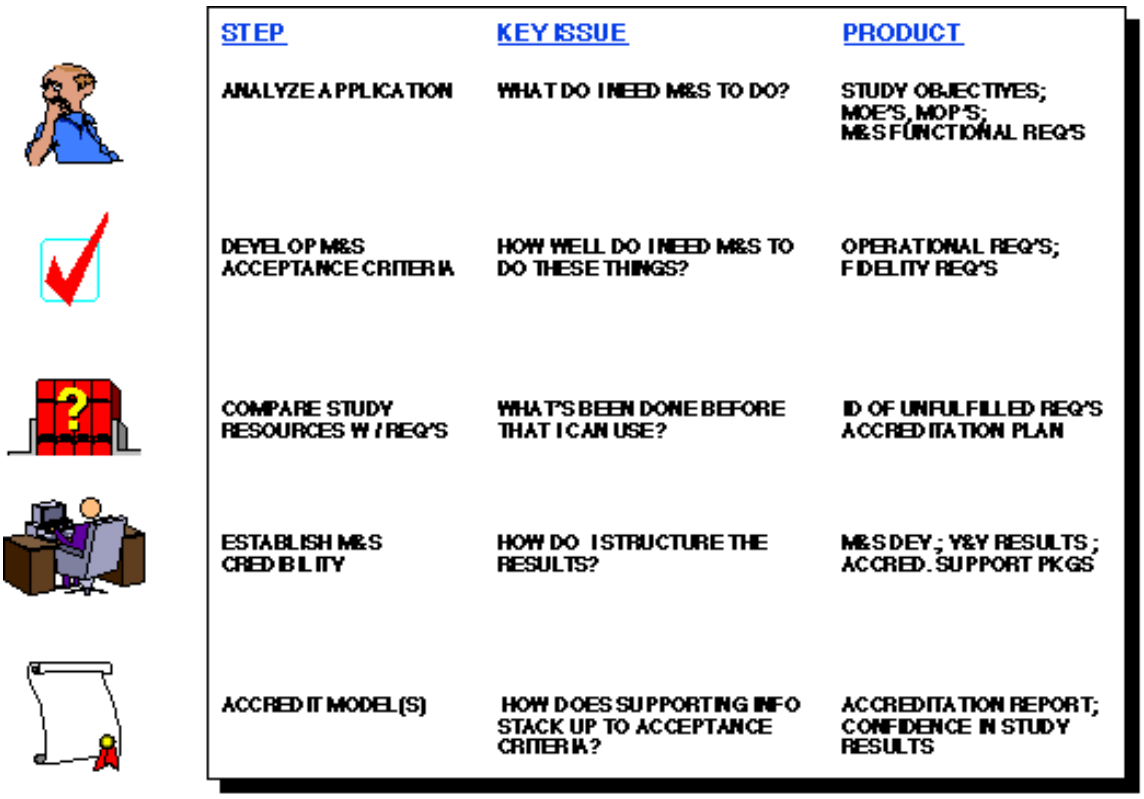

Figure 5: Relating "V&V" to "A"

The functional, operational and fidelity requirements developed in the first two steps constitute the basic checklists of acceptance criteria against which M&S characteristics will be compared when it is time to make the accreditation decision. This comparison is the third step in accreditation support. Information on M&S capabilities is compiled from documentation, product literature and other sources, and compared to the functional requirements list to determine if any of the required functions are not modeled. Information on M&S operational characteristics (e.g., how much memory it uses, what programming language it is written in, how long it takes to run a typical case, what hardware and operating system is required, what special training and maintenance is required, etc.) is obtained from similar sources, and compared to the operational requirements list to determine if additional resources will be required to maintain and operate candidate M&S during the application. Finally, the fidelity requirements list is compared to the VV&A histories and current results of the candidate M&S to determine the applicability of prior V&V, and to identify requirements for additional V&V to address the current problem. Having a history of prior V&V and accreditations is another essential factor in keeping V&V costs down.

(This latter suggests the need not only for a VV&A repository for M&S, it requires that V&V be conducted using techniques and definitions accepted by the various M&S communities, and that V&V be documented in a standardized format within each community. SMART has developed a repository for VV&A information for aircraft survivability M&S, and is in the process of transitioning it to the DoD community via the Defense Modeling and Simulation Office

(DMSO).)

Having identified any gaps in the functional, operational and fidelity requirements for candidate M&S, a final selection of M&S for the study problem is made. The acceptance criteria gaps are used to develop an accreditation plan that prioritizes each gap and describes how each will be addressed with V&V or other methods. Cost and schedule projections are made, and risk assessment and mitigation strategies are developed. The plan is briefed to the study agent (the guy who pays the bills) for approval, who also determines the frequency and format of progress reports.

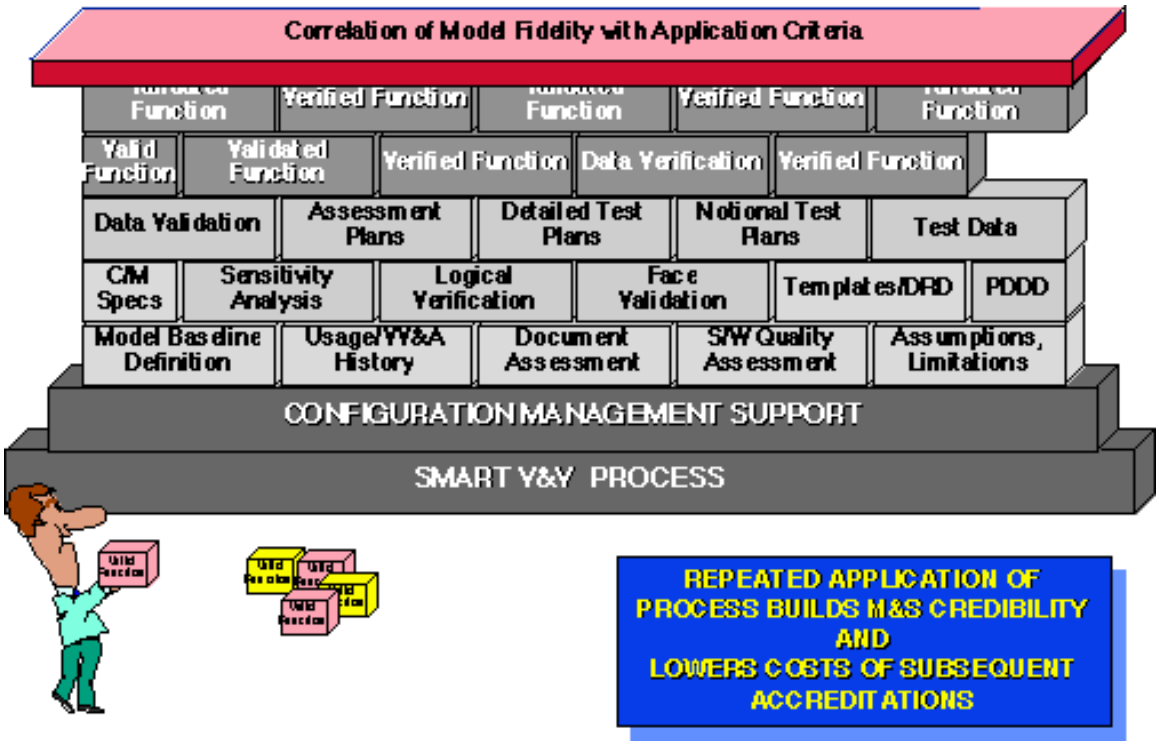

Figure 6: Building M&S Credibility

The fourth step in accreditation support process is the execution of the approved Accreditation Plan to fill the identified M&S acceptance criteria gaps. If the plan includes requirements for additional V&V, the SMART (or any other acceptable) process and products can be used to meet the fidelity requirements and report the results. If the plan includes requirements for model development, V&V should be performed concurrently. When sufficient evidence for M&S accreditation has been accumulated, (or time and/or money run out) an accreditation report is written, summarizing how the gaps in M&S acceptance criteria have been addressed and/or resolved, and providing a risk assessment of any remaining gaps. The accreditation report concludes with a recommendation to either accredit the M&S for the application at hand, or to reject it on the basis of failure to meet critical acceptance criteria. This is the fifth and final step in the accreditation support process.

The result of this five-step process is an audit trail of well defined study objectives, M&S acceptance criteria, and study results, including decisions that substantiate the use and acceptance of M&S results in the study problem. The advantages of this approach to the individual study agent, who must wrestle with an increasingly restive and skeptical decision-making community, are clear. Time spent early in the accreditation process to define the application and minimize V&V requirements in an objectively justifiable way pays dividends in reduced V&V cost and time, and in improved confidence in decisions made with M&S. Less obvious, however, is the benefit to the wider user community of the accumulation of M&S credibility in accordance with this process.

As individual models are used again and again in support of different applications, substantive, objective evidence of their credibility grows. Although each user (or group of users) contributes only those aspects of credibility applicable to the current problem, over time an objective body of evidence grows (see Figure 6). SMART's standardized V&V process, coupled with its standard V&V reporting format (the three-volume ASP format), provides a convenient way to accumulate evidence of M&S credibility. It is the steady accumulation of such evidence that reduces the cost of subsequent accreditation for individual M&S, because accreditation builds on prior V&V, rather than on independent (mostly duplicative) efforts. Seen in this light, SMART's accreditation support approach provides the DoD community a pathway to lower cost accreditation.

SMART has always been charged with the task of transferring its V&V process and products, as well as its accreditation support experience, to the wider DoD community. Ongoing efforts at applying SMART's expertise to current problems, at continuing and expanding accreditation support services to other classes of M&S, and at integrating our V&V concepts into more broadly applicable V&V guidelines are currently being funded by DMSO. We feel strongly that the incremental approach to V&V, dividing it not only into large phases, but into smaller model-level and function-level activities, reduces the tendency to "V&V for it's own sake" and helps ensure that precious V&V dollars buy meaningful V&V products.

ABOUT THE AUTHOR

The author received his B.S. in Chemistry in 1981 from St. Joseph's University in Philadelphia, and was awarded the doctorate in Physical Chemistry from Brown University in 1987. He began his career as a defense analyst at the Center for Naval Analyses in Alexandria, VA, working on technical feasibility assessments of advanced technology aircraft using M&S. While there he coordinated two separate validation efforts for the Advanced Low Altitude Radar Model (ALARM). He later became CNA's field representative to the Naval Strike Warfare Center in Fallon NV, contributing to the training syllabus in strike warfare, conducting tactical analyses and coordinating Tactics Development and Evaluation projects.

Dr. Muessig joined the Naval Air Warfare Center, Weapons Division at China Lake, CA in 1989 (then the Naval Weapons Center), working on a project to integrate data collected from strike training exercises to validate survivability assessment methodologies. It was during this time that the idea for SMART began to take shape. A proposal to the Office of the Secretary of Defense (OSD) to develop and test an integrated M&S credibility assessment process utilizing field test

data was developed, approved and funded in FY92. Since that time, Dr. Muessig has acted as Deputy Project Manager and Technical Director for the SMART Project. He is the author of numerous technical publications, most dealing with the validation of survivability M&S using test data, V&V process development, and V&V application strategies to support M&S accreditation. He is also editor of the SMART Project newsletter, SMARTALK.